DOCUMENT RESUME

ED 201 307 PS 012 127

AUTHOR Low, Lana; Wolfle, Lee M.
TITLE Enduring Effects of First-Grade Teachers on Achievement.
PUB DATE Apr 81
NOTE 26p.; Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *Elementary School Students; Grade 1; Identification; Longitudinal Studies; Measures (Individuals); *Outcomes of Education; Primary Education; *Teacher Effectiveness; *Teacher Influence
IDENTIFIERS *SRA Achievement Series

ABSTRACT

The long-term influence of effective first-grade teachers on their students was examined in this study. Subjects were about 200 students enrolled in the classes of nine first-grade teachers in 1974-1975. During that year and each of the four succeeding years, the students were administered SRA achievement tests in the fall and in the spring. One first-grade teacher (teacher #8) was identified as especially effective because of the superior achievement gains of her students in the spring of that year. Two analytical procedures were used to examine whether a subsample of 14 students of teacher #8 would continue to demonstrate superior achievements in comparison with matched samples from the other classes. The results indicate that the students of the effective first-grade teacher not only failed to show these superior performance in subsequent elementary grades but performed worse than the students of the other eight first-grade teachers. Several possible explanations for this unexpected result are offered. (Author/JA)

ENDURING EFFECTS OF FIRST-GRADE TEACHERS ON ACHIEVEMENT

Lana Low

Clinch Valley College

and

Lee M. Wolfle

Virginia Polytechnic Institute and State University

This paper was prepared for delivery at the annual meetings of the American Educational Research Association, Los Angeles, April 13-17, 1981. The authors wish to express their gratitude to Gary D Fenstermacher, Jimmie C. Fortune, Lawrence McCluskey, James Michaels, and Jerome Niles for their helpful criticisms and guidance.

ENDURING EFFECTS OF FIRST-GRADE TEACHERS ON ACHIEVEMENT

ABSTRACT

Pedersen, Faucher and Eaton (1978) concluded that first-grade teachers have enduring effects on their pupils. The purpose of the present study was to test the robustness of their finding. Like Pedersen, we believed that students who get a good start in school will benefit throughout their school careers. Much to our surprise, our results indicated that the students of one effective first-grade teacher not only failed to continue their superior performance in subsequent elementary grades, they did relatively worse than the students of other first-grade teachers. Several possibilities for this unexpected result are offered.

ENDURING EFFECTS OF FIRST-GRADE TEACHERS ON ACHIEVEMENT

Most of the existing research on teacher effectiveness concentrates on the short-term relationships between measures of teacher classroom behaviors (processes) and measures of student learning outcomes (products). Scholarly reviews of these research findings can be found in Rosenshine (1971), Rosenshine and Furst (1971), Dunkin and Biddle (1974), Brophy and Evertson (1976), Medley (1977), Borich and Finton (1977), and Fisher, et al. (1978). But there exists also the obligation to devote serious work and attention to the long-term impact of the educational enterprise. Unfortunately, conceptual, methodological, and logistic problems have hampered empirical study of the long-term effects of education (Härnqvist, 1977). As a result, there is a paucity of research on the enduring effects of education. The purpose of this article is to report on the long-term effects exhibited by students who were exposed to effective first-grade teachers.

This study takes as its point of reference the work of Pedersen, Faucher and Eaton (1978). Their remarkable study sought originally to examine atypical IQ changes, but came to focus on the enduring effects of one remarkable first-grade teacher. The first stage of their study, which was based on data from school records, sought to determine the effects of background characteristics and school experiences on changes in IQ. They found that "the subsequent pattern of IQ change does significantly vary with first-grade teacher" (Pedersen, et al., 1978, p. 6). Specifically, they found that students of one first-grade teacher, identified as Miss A, were more likely to exhibit IQ increases than students of other teachers.

The authors then extended their study to test their belief that if pupils get a good start in school, they will benefit not only throughout the school careers but throughout their adult lives as well (Pedersen, et al., 1978, p. 10). In brief, they found that Miss A's students achieved higher adult statuses than their peers in other first-grade classes, and the differential effects remained even after controlling for differences in background characteristics. Evidence suggested that a good first-grade teacher can provide children with more or head start, which will continue to be felt in later life, both from the remainder of their schooling and in their adult careers.

Nevertheless, their conclusions were based on one teacher, one school, and one set of students. The possibility remains that what we generally believe to be enduring effects of good first-grade teachers have been based on a unique instance. The purpose of this study was therefore to test the robustness of Pedersen, Faucher and Eaton's (1978) conclusions. Specifically, using achievement gains of pupils in a number of first-grade classes, we identified one good teacher whose students exhibited larger gains in achievement than did students of other first grade teachers. Having identified one good first-grade teacher on the basis of superior achievement of her students, we then asked whether the students continued to demonstrate superior achievements in subsequent elementary grades.

THE DATA

Our data were drawn from students of nine first-grade teachers in a rural county in Virginia. The pupils included in the analysis were those

who were enrolled in first grade in 1974-75. During that school year, and each year thereafter, the pupils were administered SRA achievement tests twice each year -- once in the fall shortly after the school year began, and again in the spring shortly before it closed. These data were taken from school records. In addition, we recorded the student's age in months, the student's gender, and developed an admittedly crude index of socio-economic status by noting which students were eligible for a free lunch program.

The SRA test battery, which forms the central core of our analysis, consisted of the primary and multilevel editions of the SRA achievement series. Within the primary edition, Primary I was administered for the fall and spring testing of grade one, and the fall testing of grade two; Primary II was used for the spring testing in grade two, and the fall and spring testing of grade three. The multilevel edition (blue level) was used in grades four and five. The test scores used in the analysis were growth scale values; these are numeric scales which provide continuous measurement of progress from grade one to grade twelve. The scale ranges from zero to about 850, and there is a growth scale value equivalent to each raw score from every level of the SRA achievement series.

When the 1974 cohort of pupils entered the first grade, there were 195 students enrolled in the nine classrooms scattered across the county. We believe this data set is particularly appropriate for our analysis, because (unlike many urban school systems) there was a low rate of attrition of the student sample; of the 195 students in the initial first-grade sample, there were 134 on whom achievement data were available for the entire five year period of study.

IDENTIFYING EFFECTIVE FIRST-GRADE TEACHERS

In order to determine whether students of effective first-grade teachers continue to demonstrate superior achievements in subsequent elementary grades, it was first necessary to identify one or more effective first-grade teachers. We accomplished this task by comparing the first-grade spring achievement scores to scores predicted on the basis of each student's fall 1974 test score, their sex, age, and socioeconomic status. The difference between each student's actual spring test score and their predicted score was calculated. Mean residualized gain scores were then computed for each teacher. These are shown in Table 1. The table clearly shows that the students of teacher #8 achieved spring first-grade scores far in excess of their predicted scores. The mean score of 54.5 is significantly different at the .05 level from those of every other teacher.

We want to interpret this finding as indicative of teacher #8's effectiveness as a first-grade teacher. We have not yet considered, however, some possible alternative explanations. It is sometimes possible to achieve a large residualized gain by the artifact of starting at a low level of achievement, and merely catching up. The gains exhibited by the students of teacher #8, however, were not due to regression toward the mean. When we examined the fall first-grade SRA scores, we did find one teacher -- teacher #3 -- whose students had significantly higher scores than those of the other eight teachers, but none of the latter's scores were significantly different from each other. Most of the first-grade teachers began with students of equal ability, but at the end of first grade

Table 1. One-Way Analysis of Variance with Teacher Means and Standard Deviations for First Grade: 1974-75

Teacher	Number of Students	Mean Residualized Gain	Standard Deviation
1	24	-17.40	40.58
2	21	11.99	48.26
3	24	-15.46	27.69
4	19	7.18	35.88
5	26	-24.76	39.53
6	21	26.81	29.84
7	27	- 8.94	35.89
8	19	54.54	32.52
9	14	- 2.93	34.08

F-Ratio	Probability
9.943	< 0.0001

teacher #8's students were outscoring their peers in other classrooms. Another possibility exists. It is possible we misidentified teacher #8 as an effective teacher because of some unique trait of the particular 1974-75 first-grade cohort. Would she be as effective with other classes? To answer this question we examined the fall to spring first-grade gains in two additional years, 1975-76 and 1977-78, for the identically same nine first-grade teachers. In both of these additional years, the students of teacher #8 again made gains in achievement greater than those of the other eight teachers. In comparison with teacher #8, in 1974-75 the second ranked teacher was teacher #6, but in 1975-76 she was ranked fourth, and in 1977-78 she was ranked seventh. In short, the students of teacher #8 began first grade indistinguishable from students of other teachers, but ended the year with significantly larger achievement scores. Indeed, teacher #8's first-grade students performed similarly year after year. This pattern is unlike that of any other teacher in the sample. We believe we have identified our effective teacher. The next step was to see how the 1974-75 students of this teacher performed in subsequent grades.

METHODS OF MEASURING ENDURING EFFECTS

Having identified an effective first-grade teacher, we then addressed the question of whether exposure to teacher #8 in the first grade produced enduring effects which could be measured in terms of superior achievement in subsequent elementary grades. Unfortunately, no uniformly satisfactory methodological procedure exists to unambiguously answer the question. Without such a procedure, the approach that was adopted was to use two

distinctly different methodological procedures. If the conclusions from each procedure match in terms of their substantive interpretations, one can be more confident that the results depended upon the presence or absence of enduring effects, and were not dependent upon the assumptions of a particular methodology.

The first approach taken was to consider the students of teacher #8 to be a subset of a population represented by students of the other eight first-grade teachers. As discussed above, there was some evidence to believe that at least prior to entering first grade, the students of teacher #8 were indistinguishable from the students of the other eight teachers. By considering the students of teacher #8 initially representative of a population of all students, the mean levels of achievement of teacher #8's students were compared to those of the other teachers. Had there been only one pair of means to compare, a univariate t-test would have been appropriate; but there were four pairs of means to compare, one for each of the subsequent elementary grades. Accordingly, a Hotelling's T^2 statistic was used to determine whether the set of achievement scores of the students of teacher #8 measured in the spring of grades two, three, four, and five, were significantly different from a set of achievement scores assumed to be population means; i.e., the scores of the students of all other teachers. The scores used in this analysis were spring SRA achievement scores, not gain scores. We expected to find that the level of achievement of teacher #8's students would exceed those of the other eight teachers' students.

The second approach taken to measure the enduring effects of teacher #8 was to match students of teacher #8 with comparable students of the

other eight teachers. This approach arose from the concern that the students of teacher #8 performed at a higher level of achievement in first grade because they were intellectually superior to the students of the other eight first-grade teachers. While there is no evidence to suggest that teacher #8's students began first grade with higher levels of achievement, they may have possessed a greater propensity for achievement. If so, teacher #8 has been singled out not for her teaching, but because of the characteristics of her students. Accordingly, a procedure suggested by Rubin (1973) was adopted for this study. Rubin suggested matched sampling as a method of reducing bias and increasing precision in observational studies, especially when random assignment of treatments to subjects was absent. In this study, the 14 students of teacher #8 were matched with an equal number of students from a pool representing the other eight first-grade teachers. The students were matched individually on the basis of spring 1975 achievement scores. Once the 14 matches were complete, the 28 students were used to estimate the effects of exposure to teacher #8 versus the eight other first-grade teachers across four subsequent years of elementary school. A split-plot analysis of variance design was used to estimate the teacher effect, the grade effect, and the teacher-grade interaction. Significant results on this analysis, on the Hotelling's T^2 analysis, or preferably both, would indicate that teacher #8's students demonstrated levels of achievement superior to those of students of the other eight teachers.

THE ENDURING EFFECTS OF TEACHER #8

Using the two analytical procedures outlined in the previous section, we proceeded to the question of whether students exposed to an effective

first-grade teacher continue to demonstrate superior achievements in subsequent elementary grades. This was our expectation. It was not what we found.

We first compared the mean levels of achievement (spring SRA scores) of teacher #8's students in grades two through five against the scores of all other students. These results are shown in Table 2. The Hotelling's T^2 statistic was statistically significant; but contrary to our expectations, the mean achievement scores of teacher #8's students were generally smaller than those of the other students. The construction of simultaneous confidence intervals about the sample means revealed at the .05 level that the means for fourth and fifth grade were significantly different. By fourth grade, the students of the eight less effective first-grade teachers were scoring higher on the SRA schievement tests than were students of teacher #8.

A second analytical strategy was used to determine the enduring effects of teacher #8. In this instance, we matched 14 students of teacher #8 with 14 students of the other eight teachers camparable on the basis of spring 1975 scores (end of first grade). These data were analyzed by employing a split-plot analysis of variance. The mean growth scale values are shown in Table 3, which also shows the F-values associated with tests of teacher effects, grade effects, and teacher-grade interactions. These indicate all three comparisons are significant. The mean scores vary by grade. They vary between the 14 students of teacher #8 and the 14 matched peers from other classes. The significant interaction indicates the teacher differences are not constant from grade to grade. In substantive terms, while the mean achievement of the two sets of students were

Table 2. Means and Standard Deviations of 14 Students of Teacher #8 as Compared with Means of All Other Students (N=120): Grades Two, Three, Four and Five

Grade	All Other Students: Mean	Teacher #8	
		Mean *	Std. Dev.
Two	192.6	210.3	47.8
Three	245.7	225.7	54.2
Four	274.2	231.0	57.7
Five	312.4	273.1	53.7

* Hotelling's T^2 = 25.27

Table 3. Mean Spring SRA Achievement Scores by First Grade Teacher and Elementary Grades Two Through Five

	Grade 2	Grade 3	Grade 4	Grade 5	Mean Total
Teacher #8	210.3	225.7	231.0	273.1	235.0
Matched Sample	232.3	294.7	327.9	379.1	308.5
Total	221.3	260.2	279.4	326.1	

Teacher: F(df: 1,26) = 19.14
Grade: F(df: 3,78) = 70.52
Interaction: F(df: 3,78) = 13.23

equal at the end of first grade, by the end of second grade the 14 students of teacher #8 were nearly 20 points below the average of a matched sample. By the end of third grade, the difference was 70 points, and continued to grow in grades four and five.

To our surprise, the results of these two analytical procedures indicate that the students of the effective first-grade teacher not only failed to continue their superior performance in subsequent elementary grades, they did relatively worse than the students of the other eight first-grade teachers. This was true whether compared to all other students, or a matched sample.

DISCUSSION

We began our investigation with the intention of replicating Pedersen, Faucher and Eaton's (1978) conclusions that students who get a good start in school will benefit throughout their school careers. During the first phase of our study, we identified an effective first-grade teacher on the basis of the short-term gains in achievement exhibited by her students. This analysis confirmed our anecdotal impressions about teacher #8: her students spent more time on task than students of other teachers; she was respected by her former students and students' parents; her supervisors held her in high esteem. Thus, it was with some surprise that when we turned our attention away from short-term gains, and focused instead on long-term differences in achievement, the students of our effective teacher did not perform as well as students of other teachers. These unexpected results demanded explanation. Unfortunately, they were truly unexpected, and the present study was not designed to answer

satisfactorily the question of why initial gains made in first grade were not sustained. A lot of what follows is just informed conjecture.

We first considered whether socioeconomic differences could account for why teacher #8's students performed below the level of their matched peers. Unfortunately, the only measure of socioeconomic status available was whether or not a student received free lunches, which merely distinguished the poor from the nonpoor. We used this fallible measure in a re-estimation of the split-plot analysis of variance. It had no effect -- there were no socioeconomic main effects, no interactions between socioeconomic status and other variables, and the teacher effect and teacher-grade interactions remained statistically significant. Whatever produced the differences in long-term achievement, it was apparently not due to socioeconomic differences.

We next considered whether teacher #8's students were disproportionately assigned to poorer teachers in subsequent grades. This question seems crucial, because if true it would imply that the subsequent lower levels of achievement of teacher #8's students were not attributable to her. But once again the evidence does not exist to indicate teacher #8's students were directed to poorer teachers in subsequent elementary grades. First, we had the testimony of the county's school superintendent that the teachers in teacher #8's school were no better or worse than teachers in the rest of the county. Second, the school superintendent made available to us the second-grade teacher assignments of the 195 students in our analysis, and these data were used to measure residualized gain scores across second-grade teachers. If the students of teacher #8 in the first grade had been assigned to second-grade teachers whose students exhibited

lower than average gains, then one might conclude that lower performance in subsequent grades was due to second-grade teachers. If, however, the students of teacher #8 in the first grade had been assigned to second-grade teachers whose students exhibited average gains or higher than average gains, then one could not conclude that the former students of teacher #8 had been assigned disproportionately to poorer teachers in the second grade.

To effect this analysis, the spring 1976 achievement scores were regressed on the fall 1975 achievement scores, age, socioeconomic status, and sex. The residuals were then analyzed in a one-way analysis of variance across 15 second-grade teachers (see Table 4). Teacher #8 had 16 students represented in the analysis; 6 students were assigned to second-grade teacher M, and 10 students were assigned to second-grade teacher N. Although teacher M's students had the second lowest mean gain compared to the other 14 second-grade teachers, a post hoc Tukey range test showed that the mean achievement of teacher M's students was not significantly different from 12 of the other teachers' mean achievement gains. The mean achievement gain of those 10 students assigned to teacher N had the second highest mean achievement gain of all the other fourteen second-grade teachers in this study. On the basis of these results, one is led to believe that the former students of teacher #8 did not perform at a lower level in subsequent grades because they had been disproportionately assigned to poorer teachers in second grade. One second-grade teacher (to whom was assigned teacher #8's students) was clearly above average in producing achievement gains, while the other was indistinguishable from most other second-grade teachers.

Table 4. One-Way Analysis of Variance with Teacher Means and Standard Deviations for Second Grade

First Grade Teacher	Second Grade Teacher	Number of Students	Mean Residualized Gain	Standard Deviation
1	A	11	- 8.56	16.69
1	B	7	72.27	22.61
2	C	21	11.64	25.13
3	D	21	-21.08	26.61
4	E	18	- .72	16.38
5	F	23	- 6.39	28.93
6	G	5	15.28	16.33
6	H	9	.71	41.42
6	I	5	23.89	20.29
7	J	7	-37.04	42.96
7	K	10	-13.48	13.88
7	L	8	-19.77	17.89
8	M	6	-24.27	20.41
8	N	10	58.16	14.11
9	O	10	-14.73	25.65

F-Ratio	Probability
12.4860	<0.0001

If teacher #8's students had large achievement gains in first grade and gains in second grade equivalent to most other teachers, how could he students end second grade with lower scores? This problem led us to examine achievement gains and losses over the summer months. Table 5 shows the mean growth scale values for the nine first-grade teachers for all five years of the study. Comparison of the spring 1975 and fall 1975 scores shows that five of the nine classes exhibited losses over the summer, but the decline of teacher #8's students exceeded by a large margin the losses of any other teachers' students. When we refocused our interests from achivement gains exhibited during the first-grade school year to the longer term, the gains made by teacher #8's students were not even sustained over the summer months.

We have postponed until now our consideration of whether the effects we have discovered could be attributed to the instructional practices of teacher #8. In one sense, these would be the most intuitively pleasing explanations, but we first wanted to dismiss as best we could some artifactual explanations. Our consideration of the pedagogical style of teacher #8 began with an apparently simple question: Was it possible that teacher #8 was teaching to the SRA tests? We do not mean this critically. Several years ago (about the time the cohort of pupils in this research were in first grade) the opportunity arose to visit teacher #8, and to observe her in her classroom. These observations revealed that she emphasized basic skills, particularly reading and arithmetic. That she was successful is evidenced by her students' performance on the spring 1975 SRA tests. Indeed, it was the success of her students on the first-grade test which led us to single out teacher #8 as an effective teacher.

Table 5. Mean Fall and Spring Achievement Scores by First-Grade Teachers for 1974-79

Teacher	Grade One 1974-75 Fall	Grade One 1974-75 Spring	Grade Two 1975-76 Fall	Grade Two 1975-76 Spring	Grade Three 1976-77 Fall	Grade Three 1976-77 Spring	Grade Four 1977-78 Fall	Grade Four 1977-78 Spring	Grade Five 1978-79 Fall	Grade Five 1978-79 Spring
1	54.13	110.87	120.00	213.79	171.39	227.06	228.72	263.94	271.00	308.15
2	38.76	126.71	137.62	213.05	190.89	239.89	236.53	273.05	276.65	315.40
3	97.46	156.21	136.86	179.81	197.21	250.16	242.94	294.22	277.05	316.74
4	54.74	122.63	91.74	165.94	176.40	256.64	237.69	241.30	264.36	315.50
5	45.35	96.15	98.70	166.71	181.15	231.33	242.12	281.72	287.94	336.53
6	43.00	146.09	135.10	211.00	218.15	283.15	267.00	268.44	265.82	321.67
7	45.61	112.14	122.07	168.92	177.77	229.83	222.04	264.74	252.00	273.96
8	46.52	176.53	120.53	204.76	181.23	230.19	218.71	231.00	252.07	285.00
9	54.36	127.28	126.73	178.27	182.36	247.37	239.11	274.20	283.37	324.75

When we refocused our attention to longer-term achievements, however, teacher #8's students no longer excelled. The reason may be that the pedagogical style required for short-term success is not the pedagogical style that produces long-term gains. Teacher #8's emphasis on task-oriented skills was apparently not the preparation necessary for success in later elementary grades. Given this explanation, perhaps we should have anticipated the empirical results. Fenstermacher (1979, p. 174), for example, has suggested that the way basic skills are handled in schools today may impair the student's ability to pursue advanced knowledge.

Moreover, our results may have been anticipated by some of the Follow Through analyses. Ferb, et al. (1977) focused on the effects of several developmental curricula, and found effects similar to those observed here. The basic-skills theorists argued that rapid development of basic skills provides enduring effects; the developmentalists argued that slow, gradual, but thorough instruction through early years, produces better effectiveness yields in later years.

The data from Project Follow Through suggest that programs which emphasize academic activities produce a more or less even rate of academic progress across grades. Programs with affective emphasis, which stress self-concept and attitudes toward school, often do so at the expense of explicit academic activities. Children in such programs show less academic progress early in the program relative to children in basic-skills programs, but then catch up to their peers at a later point. Likewise, the expected pattern of performance for children in programs emphasizing problem solving and thinking, is similar to programs with affective emphasis; little

if any academic progress early in the program, followed by more substantial progress when the child can generalize his or her reasoning skills (Ferb, et al., 1977, p. 3).

From these two studies (the present one and Follow Through), neither of which was designed to test the enduring effects of basic achievement styles of teaching, one would conclude that basic skills teaching may produce short-term gains, but may eventually prove to be a detriment to learning. If these results are to be believed, the back to basics pedagogy may prove to be an educational handicap to those children exposed to its effects. These results and their implications, even if only partly true, imply that implementation of basic achievement styles of teaching should be re-examined; the long-term effects may be the opposite of those desired. In any event, more research on the subject is clearly warranted.

Another trend disturbs us. In the name of accountability, a number of school districts are measuring the effectiveness of their teachers by the achievements of their students. As a consequence, teachers are altering their teaching styles, and teaching to the tests that measure their students' performance. But if the results of our research are true, the implementation of such models of accountability may prove to be a long-term handicap to the achievement of students. Once again, we remain unsure about whether these conclusions are warranted, but the implications are so serious that they clearly deserve further study, both policy and empirical.

In the end, one simply hesitates to accept the major finding of this study: the students of effective first-grade teachers are handicapped in subsequent elementary grades. Not only is the finding counterintuitive, it contradicts most of the extant research on teacher effectiveness. Yet almost all such research has been designed to study the short-term effects of teachers. A final observation to be made in this paper therefore pertains to the nature of teacher-effectiveness research. The identification of teacher #8 was based on residualized gain scores measured from fall to spring in a single school year. Indeed, this is the common practice in short-term studies. However, the performance decline of teacher #8's students over the summer calls into question the seemingly superior achievement gains. In fact, this observed decline calls into question the current process by which short-term teacher effects are determined. If instead of using fall to spring gains, this study had used fall to fall gains, teacher #8 would not have been identified as the single most effective teacher. Perhaps if the bank of knowledge we have accumulated about teacher effectiveness had also been based on fall to fall gains, rather than fall to spring gains, the substantive deposits in the knowledge bank would have been different than those we have become accustomed to. It is a possibility worth considering.

REFERENCES

Borich, G., and Finton, K. The Appraisal of Teching: Concepts and Process. Reading, Mass.: Addison-Wesley, 1977.

Brophy, J., and Evertson, C. Learning from Teaching. Boston: Allyn and Bacon, 1976.

Duncan, M., and Biddle, B. The Study of Teaching. New York: Holt, Rinehart and Winston.

Fenstermacher G. D. A philosophical consideration of recent research on teacher effectiveness. In L. S. Shulman (ed.), Review of Research in Education, Vol. VI. Itaska, Ill.: Peacock, 1979.

Ferb, T. E., Larson, J. C., Napior, D., Mingolla, E., and Frey, P. E. Education as Experimentation: A Planned Variation Model, Vol. IV-D of A Longitudinal Study of Follow Through. Cambridge, Mass.: Abt Associates.

Härnqvist, K. Enduring effects of schooling -- a neglected area in educational research. Educational Researcher, 1977, 6 (November), 5 - 11.

Medley, D. Teacher Competence and Teacher Effectiveness. Washington, D.C.: American Association of Colleges for Teacher Education.

Pedersen, E., Faucher, T., and Eaton, W. A new perspective on the effects of first-grade teachers on children's subsequent adult status. Harvard Educational Review, 1978, 48, 1 - 31.

Rosenshine, B. Teaching Behaviors and Student Achievement. London: International Association for the Evaluation of Educational Achievement, 1971.

Rosenshine, B., and Furst, N. Research on teacher performance criteria. In B. O. Smith (ed.), Research in Teacher Education: A Symposium. Englewood Cliffs, N.J.: Prentice Hall.

Rubin, D. Matching to remove bias in observational studies. Biometrics, 1973, 29, 159 - 183.